arXiv:1112.5847v1 [math.PR] 26 Dec 2011

A LOWER BOUND FOR THE MIXING TIME OF THE RANDOM-TO-RANDOM INSERTIONS SHUFFLE

By Eliran Subag*

Technion

The best known lower and upper bounds on the mixing time for the random-to-random insertions shuffle are $\left(\frac{1}{2}-o(1)\right)n\log n$ and $(2+o(1))\,n\log n$. A long standing open problem is to prove that the mixing time exhibits a cutoff. In particular, Diaconis (2003) conjectured that the cutoff occurs at $\frac{3}{4}n\log n$. Our main result is a lower bound of $t_n=\left(\frac{3}{4}-o(1)\right)n\log n$, corresponding to this conjecture.

Our method is based on analysis of the positions of cards yet-to-be-removed. We show that for large n and t_n as above, with high probability the number of cards within a certain distance from their initial position is the same under the measure induced by the shuffle and under the stationary measure, up to a lower order term. However, under the induced measure, this lower order term is dominated by the cards yet-to-be-removed, and is of higher order than for the stationary measure.

1. Introduction. In the random-to-random insertions shuffle a card is chosen at random, removed from the deck and reinserted in a random position. Assuming the cards are numbered from 1 to n, let us identify an ordered deck with the permutation $\sigma\in S_n$ such that $\sigma(j)$ is the position of the card numbered j. The shuffling process induces a random walk Π_t, $t=0,1,\ldots$, on S_n. Let P^n_σ be the probability measure corresponding to the random walk starting from $\sigma\in S_n$.

Clearly, Π_t is an irreducible and aperiodic Markov chain. Therefore $P^n_\sigma(\Pi_t\in\cdot)$ converges, as $t\to\infty$, to the stationary measure U^n, the uniform measure on S_n. To quantify the distance from stationarity, one usually uses the total variation distance

$$d_n(t)\triangleq\max_{\sigma\in S_n}\left\|P^n_\sigma(\Pi_t\in\cdot)-U^n\right\|_{TV}=\left\|P^n_{id}(\Pi_t\in\cdot)-U^n\right\|_{TV},$$

where equality follows since the chain is transitive. The mixing time is then defined by

$$t^{(n)}_{mix}(\varepsilon)\triangleq\min\{t:\,d_n(t)\le\varepsilon\}.$$

*Research supported in part by Israel Science Foundation 853/10 and USAFOSR FA8655-11-1-3039.

AMS 2000 subject classifications: Primary 60J10.

Keywords and phrases: Mixing-time, card shuffling, random insertions.

In order to study the rate of convergence to stationarity for large n, one studies how the mixing time grows as $n \to \infty$. In particular, one is interested in finding conditions on $(t_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty} d_n(t_n)$ equals 0 or 1.

The random-to-random insertions shuffle is known to have a pre-cutoff of order $O(n\log n)$. Namely, for $c_1 = \frac{1}{2}$, $c_2 = 2$:

(i) for any sequence of the form $t_n = c_1 n\log n - k_n n$ with $\lim_{n\to\infty} k_n = \infty$, $\lim_{n\to\infty} d_n(t_n) = 1$; and
(ii) for any sequence of the form $t_n = c_2 n\log n + k_n n$ with $\lim_{n\to\infty} k_n = \infty$, $\lim_{n\to\infty} d_n(t_n) = 0$.

In Diaconis and Saloff-Coste (1993) the mixing time is shown to be of order $O(n\log n)$. Uyemura-Reyes (2002) uses a comparison technique from Diaconis and Saloff-Coste (1993) to show that the upper bound above holds with $c_2 = 4$ and proves the lower bound with $c_1 = \frac{1}{2}$ by studying the longest increasing subsequence. In Saloff-Coste and Zúñiga (2008) the upper bound is improved, also by applying a comparison technique, and shown to hold with $c_2 = 2$. An alternative proof to the lower bound with $c_1 = \frac{1}{2}$ is also given there.

A long standing open problem is to prove the existence of a cutoff in total variation (see Diaconis and Saloff-Coste (1995); Diaconis (2003)); that is, a value c such that for any $\varepsilon > 0$:

(i) for any sequence $t_n \leq (c-\varepsilon)\, n\log n$, $\lim_{n\to\infty} d_n(t_n) = 1$; and
(ii) for any sequence $t_n \geq (c+\varepsilon)\, n\log n$, $\lim_{n\to\infty} d_n(t_n) = 0$.

In particular, in Diaconis (2003) it is conjectured that there is a cutoff at $\frac{3}{4} n\log n$.

Our main result is a lower bound on the mixing time with this rate.

Theorem 1.1. *Let* $t_n = \frac{3}{4} n\log n - \frac{5}{4} n\log\log n - c_n n$ *with* $\lim_{n\to\infty} c_n = \infty$. *Then* $\lim_{n\to\infty} d_n(t_n) = 1$.

The proof is based on analysis of the distribution of the positions of cards yet-to-be-removed. Let $[n] = \{1,\dots,n\}$ and denote the set of cards that have not been chosen for removal and reinsertion up to time t by $A^t = A^{n,t}$. The following result describes the limiting distribution for a card in A^t as the size of the deck grows (in the sense below).

Let $\Rightarrow$ denote weak convergence and $N(0,1)$ denote the standard normal distribution.

THEOREM 1.2. *Let $j_n \in [n]$ and $t_n \in \mathbb{N}$ be sequences. Assume that $\gamma \triangleq \lim_{n\to\infty} \frac{j_n}{n}$ exists, and that*

$$\lim_{n\to\infty} \frac{n^2}{t_n j_n (n-j_n)} = \lim_{n\to\infty} \frac{t_n}{j_n (n-j_n)} = 0.$$

Then

$$P_{id}^n \left(\frac{\Pi_{t_n}(j_n) - j_n}{\sqrt{2t_n\lambda_n}} \in \cdot \,\middle|\, j_n \in A^{t_n} \right) \Longrightarrow P\left(N(0,1) \in \cdot\right),$$

where

$$\lambda_n = \begin{cases} \frac{j_n}{n} & \textit{if } \gamma = 0, \\ \frac{n-j_n}{n} & \textit{if } \gamma = 1, \\ \gamma(1-\gamma) & \textit{if } \gamma \in (0,1). \end{cases}$$

This can be explained by the following heuristic. For $i \in [n]$ not too close to 1 or n, for $m < t$, the conditional transition probabilities (given in (2.2) below)

$$P_{id}^n \left(\Pi_{m+1}(j) = i + k \middle| \, \Pi_m(j) = i, \, j \in A^t \right),$$

are close to symmetric. Thus, under mild conditions on t and j, conditioned on $j \in A^t$, we expect $\max_{0\le m\le t} |\Pi_m(j) - j|$ to be small. On a sufficiently small neighborhood of j the transition probabilities above (which depend on i) hardly vary. Therefore $\Pi_t(j) - j$ is roughly a sum of t 'small' i.i.d random variables.

To distinguish $P_{id}^n(\Pi_{t_n} \in \cdot)$, with t_n as in Theorem 1.1, from U^n, we study the size of sets of the form

$$\triangle_\alpha(\sigma) \triangleq \left\{ j \in D^n : \, |\sigma(j) - j| \le \alpha\sqrt{n\log n} \right\} \;, \; \sigma \in S_n,$$

where $D^n = [n] \cap [n(1-\varepsilon)/2, n(1+\varepsilon)/2]$, for fixed $\varepsilon \in (0,1)$ and a parameter $\alpha > 0$. We shall see that, as long as c_n does not approach ∞ too fast,

$$|\triangle_\alpha| / \left(2\varepsilon\alpha\sqrt{n\log n}\right) \to 1 \quad \text{in probability},$$

under both measures. That is, the probability that

$$\left| |\triangle_\alpha| / \left(2\varepsilon\alpha\sqrt{n\log n}\right) - 1 \right| > \delta$$

approaches 0, as $n \to \infty$, for any $\delta > 0$. However, the deviation $|\triangle_\alpha| - \left(2\varepsilon\alpha\sqrt{n\log n}\right)$, which for $P_{id}^n(\Pi_{t_n} \in \cdot)$ is dominated by $\left|\triangle_\alpha \cap A^{t_n}\right|$, i.e. by the cards yet-to-be-removed, is of different order for the two measures.

In Section 2 we prove Theorem 1.2 and other related results. We analyze the distribution of $|\triangle_\alpha|$ under U^n, and the distributions of $\left|\triangle_\alpha \cap A^{t_n}\right|$ and $\left|\triangle_\alpha \setminus A^{t_n}\right|$ under P_{id}^n in Section 3. The proof of Theorem 1.1, given in Section 4, then easily follows. Lastly, in Section 5 we prove two results which are used in the previous sections.

2. The Position of Cards Yet-to-be-Removed. In this section we prove Theorem 1.2 and other related results.

The increment distribution of Π_t is given by

$$\mu(\tau) = \begin{cases} 1/n & \text{if } \tau = id, \\ 2/n^2 & \text{if } \tau = (i,j) \text{ with } 1 \le i,j \le n \text{ and } |i-j| = 1, \\ 1/n^2 & \text{if } \tau = c_{i,j} \text{ with } 1 \le i,j \le n \text{ and } |i-j| > 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

where $c_{i,j}$ is the cycle corresponding to removing the card in position i and reinserting it in position j, that is

$$c_{i,j} = \begin{cases} id & \text{if } i = j, \\ (j, j-1, \dots, i+1, i) & \text{if } i < j, \\ (j, j+1, \dots, i-1, i) & \text{if } i > j. \end{cases}$$

Let $2 \le n \in \mathbb{N}$ and $j \in [n]$. Under conditioning on $\left\{j \in A^t\right\}$, $\Pi_m(j)$, $m = 0, \dots, t$ is a time homogeneous Markov chain with transition probabilities

$$\begin{aligned} p_{i,i+k}^{\Pi(j)} &\triangleq P_{id}^n\left(\Pi_{m+1}(j) = i+k \middle| \Pi_m(j) = i,\, j \in A^t\right) \\ &= \begin{cases} \frac{i(n-i)}{n(n-1)} & \text{if } k = +1, \\ \frac{(i-1)(n-i+1)}{n(n-1)} & \text{if } k = -1, \\ \frac{(i-1)^2+(n-i)^2}{n(n-1)} & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{2.2}$$

One of the difficulties in analyzing the chain is the fact that the transition probabilities $p_{i,i+k}^{\Pi(j)}$ are inhomogeneous in i. To overcome this, we consider a modification of the process for which inhomogeneity is 'truncated' by setting transition probabilities far from the initial state to be identical to these in the initial state. As we shall see, a bound on the total variation distance of the marginal distributions of the modified and original processes is easily established.

For $j \in [n]$ and $M > 0$, let $\overline{j \pm M} \triangleq [n] \cap [j-M, j+M]$ and let $\zeta_m = \zeta_m^{n,j,M}$, $m = 0, 1, \ldots$ be a Markov process starting at $\zeta_0 = j$, with transition probabilities $p_{i,i+k}^{\zeta,j,M} \triangleq P\left(\zeta_{m+1} = i+k \middle| \zeta_m = i\right)$ such that

$$\forall i \in \overline{j \pm M} : p_{i,i+k}^{\zeta,j,M} = p_{i,i+k}^{\Pi(j)},$$
$$\forall i \in \mathbb{Z} \setminus \overline{j \pm M} : p_{i,i+k}^{\zeta,j,M} = p_{j,j+k}^{\Pi(j)}.$$

Clearly, for any sequence $(k_m)_{m=0}^t \in \mathbb{Z}^{t+1}$, if $\max_{0 \le m \le t} |k_m - j| \le M$ then

$$P\left((\zeta_m)_{m=0}^t = (k_m)_{m=0}^t\right) = P_{id}^n\left((\Pi_m(j))_{m=0}^t = (k_m)_{m=0}^t \middle| j \in A^t\right). \tag{2.3}$$

Therefore, by taking complements, for any $u \le M$,

$$P_{id}^n\left(\max_{0\le m\le t} |\Pi_m(j) - j| > u \middle| j \in A^t\right) = P\left(\max_{0\le m\le t} |\zeta_m - j| > u\right). \tag{2.4}$$

Moreover, (2.3) implies that for any $B \subset \mathbb{Z}^{t+1}$

$$\begin{aligned}
&P_{id}^n\left((\Pi_m(j))_{m=0}^t \in B \middle| j \in A^t\right) - P\left((\zeta_m)_{m=0}^t \in B\right) \\
&\quad = P_{id}^n\left((\Pi_m(j))_{m=0}^t \in B, \max_{0\le m\le t} |\Pi_m(j) - j| > M \middle| j \in A^t\right) \\
&\qquad - P\left((\zeta_m)_{m=0}^t \in B, \max_{0\le m\le t} |\zeta_m - j| > M\right).
\end{aligned}$$

Since both terms in the last equality are bounded from above by the equal expressions of (2.4) (and from below by zero), it follows that

$$\left\| P_{id}^n\left((\Pi_m(j))_{m=0}^t \in \cdot \middle| j \in A^t\right) - P\left((\zeta_m)_{m=0}^t \in \cdot\right) \right\|_{TV}$$

$$\le P\left(\max_{0\le m\le t} |\zeta_m - j| > M\right) \tag{2.5}$$

A simple computation shows that $\left|p_{i,i+1}^{\Pi(j)} - p_{i,i-1}^{\Pi(j)}\right|$ is bounded by $\frac{1}{n}$ for any i. On the other hand, $p_{i,i\pm1}^{\Pi(j)}$ is roughly equal to $i(n-i)/n^2$. Thus if j is large enough and M, and thus $\left|\overline{j \pm M}\right|$, is small compared to j, we can think of $\zeta_m^{n,j,M}$ as a perturbation of a random walk with a very small bias. In order to make this precise we decompose $\zeta_m^{n,j,M}$ as a sum of a random walk determined by the increment distribution in state j and two additional

random processes related to the 'defects' in symmetry and homogeneity in state.

Consider the vector-valued Markov process

$$(S_m, X_m, Y_m) = \left(S_m^{n,j,M}, X_m^{n,j,M}, Y_m^{n,j,M}\right)$$

starting at $(S_0, X_0, Y_0) = (0,0,0)$ with transition probabilities as follows. For each $k \in \mathbb{Z}$ define

$$q_k = \min\left\{p_{k,k+1}^{\zeta,j,M}, p_{k,k-1}^{\zeta,j,M}\right\}, \tag{2.6}$$

$$r_k = \max\left\{p_{k,k+1}^{\zeta,j,M}, p_{k,k-1}^{\zeta,j,M}\right\}.$$

For a state (i_1, i_2, i_3) set $i = i_1 + i_2 + i_3$ and define

$$w_i = \arg\max_{k=\pm 1}\left(p_{j+i,j+i+k}^{\zeta,j,M}\right),$$

$$z_i = \operatorname{sgn}\left(q_j - q_{j+i}\right),$$

where sgn is the sign function (the definition of sgn at zero will not matter to us). Define the transition probabilities by

$$P\left(\left(S_{m+1}, X_{m+1}, Y_{m+1}\right) = \left(i_1 + k_1, i_2 + k_2, i_3 + k_3\right) \middle| \left(S_m, X_m, Y_m\right) = \left(i_1, i_2, i_3\right)\right)$$
$$= \begin{cases} \min\left\{q_{j+i}, q_j\right\} & \text{if } \left(k_1, k_2, k_3\right) = \left(+1, 0, 0\right), \\ \min\left\{q_{j+i}, q_j\right\} & \text{if } \left(k_1, k_2, k_3\right) = \left(-1, 0, 0\right), \\ \left|q_j - q_{j+i}\right| & \text{if } \left(k_1, k_2, k_3\right) = \left(+\frac{1+z_i}{2}, -1, 0\right), \\ \left|q_j - q_{j+i}\right| & \text{if } \left(k_1, k_2, k_3\right) = \left(-\frac{1+z_i}{2}, +1, 0\right), \\ r_{j+i} - q_{j+i}, & \text{if } \left(k_1, k_2, k_3\right) = \left(0, 0, w_i\right), \\ c_i, & \text{if } \left(k_1, k_2, k_3\right) = \left(0, 0, 0\right). \end{cases}$$

where c_i is chosen such that the sum of probabilities is 1.

It is easy to verify that $(S_m + X_m + Y_m)_{m=0}^{\infty}$ is a Markov process with transition probabilities identical to those of $(\zeta_m - j)_{m=0}^{\infty}$. Therefore the two processes have the same law. It is also easy to check that S_n is a random walk with increment distribution

$$\mu(+1) = \mu(-1) = q_j \ , \ \mu(0) = 1 - 2q_j.$$

In order to study X_m and Y_m we need the following proposition.

PROPOSITION 2.1. *Let $\{A_m\}_{m=0}^{\infty}$ and $\{B_m\}_{m=0}^{\infty}$ be integer-valued random processes starting at the same point $A_0 = B_0$. Suppose that there exist $p_{ik}^A \in [0,1]$ such that for any $m \geq 0$ and $k, i, i_0, \dots, i_{m-1} \in \mathbb{Z}$ (such that the conditional probabilities are defined)*

$$\begin{aligned} p_{ik}^A &= P\left(A_{m+1} = k \middle| A_{m+1} \neq i, A_m = i\right) \\ &= P\left(A_{m+1} = k \middle| A_{m+1} \neq i, A_m = i, A_{m-1} = i_{m-1}, \dots, A_0 = i_0\right) \end{aligned}$$

and similarly for B_m with p_{ik}^B. Assume that for any $i, k \in \mathbb{Z}$, $p_{ik}^A = p_{ik}^B$. Finally, suppose that for any $m \geq 0$ and $k, i, i_0, \dots, i_{m-1}, j_0, \dots, j_{m-1} \in \mathbb{Z}$, (whenever defined)

$$\begin{aligned} &P\left(A_{m+1} \neq i \middle| A_m = i, A_{m-1} = i_{m-1}, \dots, A_0 = i_0\right) \\ &\quad \geq P\left(B_{m+1} \neq i \middle| B_m = i, B_{m-1} = j_{m-1}, \dots, B_0 = j_0\right). \end{aligned}$$

Then for any $t \in \mathbb{N}$ and $\delta > 0$

$$P\left(\max_{0 \leq m \leq t} A_m \geq \delta \right) \geq P\left(\max_{0 \leq m \leq t} B_m \geq \delta \right),$$

and

$$P\left(\max_{0 \leq m \leq t} |A_m| \geq \delta \right) \geq P\left(\max_{0 \leq m \leq t} |B_m| \geq \delta \right).$$

PROOF. The processes $\{A_m\}$ and $\{B_m\}$ can be coupled so that they jump from a given state to a new state according to the same order of states, say according to the order $\{k_m\}_{m=0}^{\infty}$, and such that the amount of time that $\{B_m\}$ spends in any given state k_m before jumping to state k_{m+1} is at least as much as $\{A_m\}$ spends there. The proposition follows easily from this. $\square$

The only nonzero increments of X_m are ± 1. Note that

$$\begin{aligned} &P\left(X_{m+1} = i_m + 1 \middle| X_{m+1} \neq i_m, \{X_p\}_{p=0}^m = \{i_p\}_{p=0}^m \right) \\ &= \sum_{(k_1,k_2) \in \mathbb{Z}^2} P\left(X_{m+1} = i_m + 1 \middle| X_{m+1} \neq i_m, \{X_p\}_{p=0}^m = \{i_p\}_{p=0}^m, \dots \right. \\ &\left. S_m = k_1, Y_m = k_2 \right) P\left(S_m = k_1, Y_m = k_2 \middle| X_{m+1} \neq i_m, \{X_p\}_{p=0}^m = \{i_p\}_{p=0}^m \right) \\ &= \frac{1}{2}. \end{aligned}$$

The last equality follows from Markov property of (S_m, X_m, Y_m). The same, of course, holds for the negative increment. In addition, again by Markov

property,

$$
\begin{aligned}
&P\left(X_{m+1}\neq i_m \middle| \left\{X_p\right\}_{p=0}^m=\left\{i_p\right\}_{p=0}^m\right)\\
&\quad=\sum_{(k_1,k_2)\in\mathbb{Z}^2} P\left(X_{m+1}\neq i_m \middle| \left\{X_p\right\}_{p=0}^m=\left\{i_p\right\}_{p=0}^m, S_m=k_1, Y_m=k_2\right)\times\\
&\quad\ P\left(S_m=k_1, Y_m=k_2 \middle| \left\{X_p\right\}_{p=0}^m=\left\{i_p\right\}_{p=0}^m\right)\\
&\quad\leq \max_{k_1,k_2} P\left(X_{m+1}\neq i_m \middle| X_m=i_m, S_m=k_1, Y_m=k_2\right)\leq 2\max_{i\in\overline{j\pm M}}\left|q_i-q_j\right|\\
&\quad\leq 2M\max_{x\in[1,n]}\left(\max\left\{\left|\frac{d}{dx}\frac{x\left(n-x\right)}{n\left(n-1\right)}\right|,\left|\frac{d}{dx}\frac{\left(x-1\right)\left(n-x+1\right)}{n\left(n-1\right)}\right|\right\}\right)\\
&\quad\leq \frac{2M}{n-1},
\end{aligned}
$$

where the maximum in the first inequality is over all k_1, k_2 such that the conditional probability is defined.

Thus, according to Proposition 2.1, for $\delta>0$,

$$
P\left(\max_{0\leq m\leq t}\left|X_m^{n,j,M}\right|\geq\delta\right)\leq P\left(\max_{0\leq m\leq t}\left|W_m^{n,M}\right|\geq\delta\right), \tag{2.7}
$$

where $W_m=W_m^{n,M}$ is a random walk starting at 0 with increment distribution

$$
\nu\left(+1\right)=\nu\left(-1\right)=\frac{M}{n-1}\ ,\ \nu\left(0\right)=1-2\frac{M}{n-1}.
$$

Similarly, for the process $\widetilde{Y}_t=\sum_{m=1}^t\left|Y_m-Y_{m-1}\right|$, whose increments are 0 and 1, we have

$$
\begin{aligned}
&P\left(\widetilde{Y}_{m+1}=i_m+1 \middle| \left\{\widetilde{Y}_p\right\}_{p=0}^m=\left\{i_p\right\}_{p=0}^m\right)\\
&\quad\leq \max_{k_1,k_2} P\left(Y_{m+1}\neq i_m \middle| Y_m=i_m, S_m=k_1, X_m=k_2\right)\\
&\quad\leq \max_{i\in\mathbb{Z}}\left(r_{j+i}-q_{j+i}\right)=\max_{i\in\overline{j\pm M}}\left|\frac{i\left(n-i\right)}{n\left(n-1\right)}-\frac{\left(i-1\right)\left(n-i+1\right)}{n\left(n-1\right)}\right|\\
&\quad=\max_{i\in\overline{j\pm M}}\left|\frac{n-2i+1}{n\left(n-1\right)}\right|\leq\frac{1}{n}.
\end{aligned}
$$

Therefore, for $\delta>0$,

$$
P\left(\max_{0\leq m\leq t}\left|Y_m^{n,j,M}\right|\geq\delta\right)\leq P\left(\widetilde{Y}_t\geq\delta\right)\leq P\left(N_t^n\geq\delta\right), \tag{2.8}
$$

where $N_t = N_t^n \sim \text{Bin}\left(t, \frac{1}{n}\right)$.

Since the increment distributions of W_m and S_m are symmetric, the classical Lévy inequality yields, for any $\delta > 0$,

$$P\left(\max_{0\leq m\leq t} |W_m| \geq \delta\right) \leq 4P\left(W_t \geq \delta\right). \tag{2.9}$$

and

$$P\left(\max_{0\leq m\leq t} |S_m| \geq \delta\right) \leq 4P\left(S_t \geq \delta\right). \tag{2.10}$$

Having established the connections between the different processes, we are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. The case where $\gamma = 1$ follows by symmetry from the case with $\gamma = 0$. Assume $\gamma \in [0, 1)$. In this case, the hypothesis in the theorem are equivalent to

$$\lim_{n\to\infty} \frac{n}{t_n j_n} = \lim_{n\to\infty} \frac{t_n}{n j_n} = 0.$$

Let $n \in \mathbb{N}$, $j \in [n]$ and $M > 0$. Based on (2.7)-(2.9) and a union bound, for $u \in \mathbb{R}$, $\delta > 0$, we have

$$\begin{aligned} &P\left(\zeta_t - j \geq u\right) \\ &\quad\leq P\left(S_t \geq u - \delta\right) + P\left(\max_{0\leq m\leq t} |X_m| \geq \frac{\delta}{2}\right) + P\left(\max_{0\leq m\leq t} |Y_m| \geq \frac{\delta}{2}\right) \\ &\quad\leq P\left(S_t \geq u - \delta\right) + 4P\left(W_t \geq \frac{\delta}{2}\right) + P\left(N_t \geq \frac{\delta}{2}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} &P\left(\zeta_t - j \geq u\right) \\ &\quad\geq P\left(S_t \geq u + \delta\right) - P\left(\max_{0\leq m\leq t} |X_m| \geq \frac{\delta}{2}\right) - P\left(\max_{0\leq m\leq t} |Y_m| \geq \frac{\delta}{2}\right) \\ &\quad\geq P\left(S_t \geq u + \delta\right) - 4P\left(W_t \geq \frac{\delta}{2}\right) - P\left(N_t \geq \frac{\delta}{2}\right). \end{aligned}$$

Assume $\frac{\delta}{2} - \frac{t}{n} > 0$. By computing moments and applying the Berry-Esseen theorem to approximate the tail probability function of S_t, and applying

Chebyshev's inequality to bound the tail probability functions of W_t and N_t, we arrive at

$$P\left(\zeta_t^{n,j,M} - j \geq u\right) \leq \Psi\left(\frac{u-\delta}{\sqrt{2tq_j}}\right) + \frac{C}{\sqrt{2tq_j}} + \frac{32Mt}{\delta^2\left(n-1\right)} + \frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2}-\frac{t}{n}\right)^2} \tag{2.11}$$

and

$$P\left(\zeta_t^{n,j,M} - j \geq u\right) \geq \Psi\left(\frac{u+\delta}{\sqrt{2tq_j}}\right) - \frac{C}{\sqrt{2tq_j}} - \frac{32Mt}{\delta^2\left(n-1\right)} - \frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2}-\frac{t}{n}\right)^2}, \tag{2.12}$$

where q_j is defined in (2.6), C is the constant from the Berry-Esseen theorem and Ψ is the tail probability function of a standard normal variable.

For two sequences of positive numbers v_n, v_n' let us denote $v_n \ll v_n'$ if and only if $\lim_{n\to\infty} v_n/v_n' = 0$. By assumption, $\sqrt{\frac{t_n j_n}{n}} \ll j_n$, therefore we can choose a sequence M_n such that $\sqrt{\frac{t_n j_n}{n}} \ll M_n \ll j_n$. Similarly, since $M_n \ll j_n$ we can set δ_n with $\sqrt{\frac{t_n M_n}{n}} \ll \delta_n \ll \sqrt{\frac{t_n j_n}{n}}$. As assumed, $\frac{t_n}{n} \ll j_n$ and $1 \ll \frac{t_n j_n}{n}$. Therefore $\frac{t_n}{n}, 1 \ll \sqrt{\frac{t_n j_n}{n}} \ll M_n$, which also implies that $\sqrt{\frac{t_n}{n}}, \frac{t_n}{n} \ll \delta_n$.

Now, let $x \in \mathbb{R}$ and set $u_n = x\sqrt{2t_n\lambda_n}$. Let us consider the inequalities derived from (2.11) and (2.12) by replacing each of the parameters by a corresponding element from the sequences above. Based on the relations established for the sequences and the assumptions on t_n and j_n it can be easily verified that, upon letting $n \to \infty$, all terms but those involving Ψ go to zero. Relying, in addition, on the fact that Ψ is continuous, it can be easily verified that

$$\lim_{n\to\infty} \Psi\left(\frac{u_n \pm \delta_n}{\sqrt{2t_n q_{j_n}}}\right) = \Psi\left(x\right).$$

Hence we conclude that

$$\lim_{n\to\infty} P\left(\zeta_{t_n}^{n,j_n,M_n} - j_n \geq u_n\right) = \Psi\left(x\right). \tag{2.13}$$

Based on (2.7)-(2.10),

$$
\begin{aligned}
&P\left(\max_{0\le m\le t}\left|\zeta_m^{n,j,M}-j\right|\ge M\right)\\
&\le P\left(\max_{0\le m\le t}|S_m|\ge M-\delta\right)+P\left(\max_{0\le m\le t}|X_m|\ge\frac{\delta}{2}\right)\\
&\quad+P\left(\max_{0\le m\le t}|Y_m|\ge\frac{\delta}{2}\right)\\
&\le 4P\left(S_t\ge M-\delta\right)+4P\left(W_t\ge\frac{\delta}{2}\right)+P\left(N_t\ge\frac{\delta}{2}\right)\\
&\le\frac{8tq_j}{\left(M-\delta\right)^2}+\frac{32Mt}{\delta^2\left(n-1\right)}+\frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2}-\frac{t}{n}\right)^2},
\end{aligned}
\tag{2.14}
$$

where the last inequality follows from Chebyshev's inequality.

As before, replace the parameters by the corresponding elements from the sequences and let $n\to\infty$. The middle and right-hand side summands of (2.14) were already shown to go to zero as $n\to\infty$. Since $\delta_n\ll\sqrt{\frac{t_nj_n}{n}}\ll M_n$ the additional term also goes to zero. Combined with (2.5) and (2.13) this gives

$$\lim_{n\to\infty}P_{id}^n\left(\Pi_{t_n}\left(j_n\right)-j_n\ge u_n\middle|\,j_n\in A^{t_n}\right)=\Psi\left(x\right),$$

which completes the proof. □

In Theorem 1.2 for each n only a single card j_n of the deck of size n is involved. The following gives a uniform bound (in initial position and in time) for the tail distributions of the difference from the initial position.

Theorem 2.1. *Let $\alpha>0$ and let t_n be a sequence such that $\lim_{n\to\infty}t_n=\lim_{n\to\infty}n^2/t_n=\infty$. Then*

$$\limsup_{n\to\infty}\max_{j\in[n]}P_{id}^n\left(\max_{0\le m\le t_n}\left|\Pi_m\left(j\right)-j\right|>\alpha\sqrt{\frac{t_n}{2}}\,\middle|\,j\in A^{t_n}\right)\le 4\Psi\left(\alpha\right).$$

Proof. Set $u_n=\alpha\sqrt{\frac{t_n}{2}}$ and $j_n=\left\lfloor\frac{n}{2}\right\rfloor$. Let $M_n\ge u_n$ and $\delta_n>0$ be

sequences to be determined below. From (2.4) it follows that

$$\begin{aligned}
&\max_{j\in[n]} P_{id}^{n}\left(\max_{0\le m\le t_n}\left|\Pi_m\left(j\right)-j\right|>u_n\middle| j\in A^{t_n}\right)\\
&=\max_{j\in[n]} P\left(\max_{0\le m\le t_n}\left|\zeta_m^{n,j,M_n}-j\right|>u_n\right)\\
&\le\max_{j\in[n]}\left\{P\left(\max_{0\le m\le t_n}\left|S_m^{n,j,M_n}\right|\ge u_n-\delta_n\right)+\right.\\
&\left.P\left(\max_{0\le m\le t_n}\left|W_m^{n,M_n}\right|\ge\frac{\delta_n}{2}\right)+P\left(N_{t_n}^{n}\ge\frac{\delta_n}{2}\right)\right\}.
\end{aligned}$$

It is easy to check that for the random walks S_m^{n,j,M_n}, $j\in[n]$, the probabilities of the nonzero increments, ±1, are maximal when $j=j_n$. Therefore according to Proposition 2.1,

$$\begin{aligned}
&\max_{j\in[n]} P_{id}^{n}\left(\max_{0\le m\le t_n}\left|\Pi_m\left(j\right)-j\right|>u_n\middle| j\in A^{t_n}\right)\\
&\le P\left(\max_{0\le m\le t_n}\left|S_m^{n,j_n,M_n}\right|\ge u_n-\delta_n\right)+\\
&P\left(\max_{0\le m\le t_n}\left|W_m^{n,M_n}\right|\ge\frac{\delta_n}{2}\right)+P\left(N_{t_n}^{n}\ge\frac{\delta_n}{2}\right).
\end{aligned}$$

From (2.9) and (2.10) we conclude that

$$\begin{aligned}
&\max_{j\in[n]} P_{id}^{n}\left(\max_{0\le m\le t_n}\left|\Pi_m\left(j\right)-j\right|>u_n\middle| j\in A^{t_n}\right)\\
&\le 4\left\{P\left(S_{t_n}^{n,j_n,M_n}\ge u_n-\delta_n\right)+P\left(W_{t_n}^{n,M_n}\ge\frac{\delta_n}{2}\right)+P\left(N_{t_n}^{n}\ge\frac{\delta_n}{2}\right)\right\}.
\end{aligned}$$

Finally, note that j_n and t_n meet the conditions of Theorem 1.2 with $\gamma=\frac{1}{2}$. Therefore, defining M_n and δ_n as in the proof of the theorem (which also implies $M_n\ge u_n$) and following the same arguments therein,

$$\begin{aligned}
&\lim_{n\to\infty}\left\{P\left(S_{t_n}^{n,j_n,M_n}\ge u_n-\delta_n\right)+P\left(W_{t_n}^{n,M_n}\ge\frac{\delta_n}{2}\right)+P\left(N_{t_n}^{n}\ge\frac{\delta_n}{2}\right)\right\}\\
&=\Psi\left(\lim_{n\to\infty}\frac{u_n-\delta_n}{\sqrt{t_n/2}}\right)=\Psi\left(\alpha\right).
\end{aligned}$$

This completes the proof. ∎

3. Cards of Distance $O(\sqrt{n \log n})$ from their Initial Position. The results of Section 2 show that the position of a card that has not been removed is fairly concentrated around the initial position. This, of course, is a rare event for each card under the uniform measure U^n. On the other hand, as the number of shuffles performed grows (i.e., as t grows), more cards are removed for reinsertion.

To distinguish the measure induced by the shuffle from the uniform measure, we consider the size of sets of the form

$$\triangle_\alpha(\sigma) \triangleq \left\{ j \in D^n : |\sigma(j) - j| \leq \alpha\sqrt{n \log n} \right\} \ , \ \sigma \in S_n,$$

where $D^n = [n] \cap [n(1-\varepsilon)/2, n(1+\varepsilon)/2]$ and $\varepsilon \in (0,1)$ is arbitrary and will be fixed throughout the proofs. Under U^n, for $i \neq j$, the events $\{i \in \triangle_\alpha\}$ and $\{j \in \triangle_\alpha\}$ are 'almost' independent, as $n \to \infty$. Therefore one should expect $|\triangle_\alpha| - E\{|\triangle_\alpha|\}$ to be of order $\left(E^{U^n}\{|\triangle_\alpha|\}\right)^{1/2}$. Under $P^n_{id}(\Pi_{t_n} \in \cdot)$, if $\left|A^{t_n}\right|$ is relatively small, it seems natural that the positions of the cards that have been removed are distributed approximately as they would under U^n. Thus, $\left|\triangle_\alpha \setminus A^{t_n}\right|$ under $P^n_{id}(\Pi_{t_n} \in \cdot)$ should be distributed roughly as $|\triangle_\alpha|$ is under U^n. By that logic, we need to choose t_n so that $\left|\triangle_\alpha \cap A^{t_n}\right|$ is larger than $\left(E^{U^n}\{|\triangle_\alpha|\}\right)^{1/2}$ with high probability. Requiring the expectation $E^n_{id}\left\{\left|\triangle_\alpha \cap A^{t_n}\right|\right\}$ to be larger than $\left(E^{U^n}\{|\triangle_\alpha|\}\right)^{1/2}$ leads us to set t_n to be $\frac{3}{4}n\log n$, up to a lower order term.

In order to prove the lower bound on the mixing time, we first prove the three lemmas below. The first treats the distribution of $|\triangle_\alpha|$ under U^n. The other two deal with $\left|\triangle_\alpha \cap A^{t_n}\right|$ and $\left|\triangle_\alpha \setminus A^{t_n}\right|$ under $P^n_{id}(\Pi_{t_n} \in \cdot)$.

Let R_j denote the event $\{j \in \triangle_\alpha\}$.

Lemma 3.1. *For any α, $k > 0$,*

$$\limsup_{n\to\infty} U^n\left(\left||\triangle_\alpha(\sigma)| - 2\varepsilon\alpha\sqrt{n\log n}\right| \geq k\sqrt{2\varepsilon\alpha}\,(n\log n)^{\frac{1}{4}}\right) \leq \frac{1}{k^2}.$$

Proof. Suppose n is large enough so that $n(1-\varepsilon)/2 \geq \alpha\sqrt{n\log n}$. Then

$$\begin{aligned} E^{U^n}\{|\triangle_\alpha(\sigma)|\} &= \sum_{j\in D^n} U^n(R_j) = |D^n| \frac{1 + 2\left\lfloor \alpha\sqrt{n\log n}\right\rfloor}{n} \\ &= 2\varepsilon\alpha\sqrt{n\log n} + O(1). \end{aligned}$$

The second moment satisfies the bound

$$\begin{aligned}E^{U^n}\left\{\left|\triangle_\alpha(\sigma)\right|^2\right\} &= \sum_{j\in D^n} U^n(R_j) + \sum_{i,j\in D^n: i\neq j} U^n(R_i\cap R_j)\\ &\le E^{U^n}\left\{\left|\triangle_\alpha(\sigma)\right|\right\} + \left|D^n\right|^2 \frac{\left(1+2\left\lfloor \alpha\sqrt{n\log n}\right\rfloor\right)^2}{n(n-1)}\\ &= E^{U^n}\left\{\left|\triangle_\alpha(\sigma)\right|\right\} + \frac{n}{n-1}\left(E^{U^n}\left\{\left|\triangle_\alpha(\sigma)\right|\right\}\right)^2,\end{aligned}$$

which implies

$$Var^{U^n}\left\{\left|\triangle_\alpha(\sigma)\right|\right\} \le 2\varepsilon\alpha\sqrt{n\log n} + 4\varepsilon^2\alpha^2\log n + O(1).$$

Applying Chebyshev's inequality and letting $n\to\infty$ yields the required result. □

Next, we consider $\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}$ under P^n_{id}. For this we shall need the following lemma whose proof is in Section 5.

Lemma 3.2. *Let $n,t\in\mathbb{N}$, let $B\subset[n]$ be a random set and let $D\subset[n]$ be a deterministic set, and suppose that for some $c>0$*

$$\min_{j\in D} P^n_{id}\left(j\in B \middle| j\in A^t\right)\ge c.$$

Then, denoting $K=E^n_{id}\left|D\cap A^t\right|$, for any $r\in(0,1)$,

$$P^n_{id}\left(\left|B\cap D\cap A^t\right|\le r\cdot E^n_{id}\left\{\left|B\cap D\cap A^t\right|\right\}\right)\le \frac{K+\left(1-c^2\right)K^2}{(1-r)^2c^2K^2}.$$

Let $R_{j,t}$ and $R^{A^c}_{j,t}$ denote the events $\{j\in\triangle_\alpha(\Pi_t)\}$ and $\{j\in\triangle_\alpha(\Pi_t)\}\cap\left\{j\notin A^t\right\}$, respectively. Let $p_{t,n}\triangleq P^n_{id}\left(j\in A^t\right)$ (which is, of course, independent of j).

Lemma 3.3. *Let $v(\alpha)=1-4\Psi\left(\alpha\sqrt{\frac{8}{3}}\right)$. Let $t_n\le\frac{3}{4}n\log n$ and suppose α satisfies $v(\alpha)>0$. Then, for any $r\in(0,1)$,*

$$\begin{aligned}&\limsup_{n\to\infty} P^n_{id}\left(\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\le rv(\alpha)\varepsilon np_{t_n,n}\right)\\ &\quad\le(1-r)^{-2}\left(v^{-2}(\alpha)-1\right).\end{aligned}$$

PROOF. With $\mathcal{S}(n,\alpha)$ defined by

$$\begin{aligned}
&\mathcal{S}(n,\alpha)\\
&\triangleq \min_{j\in D^n} P_{id}^n\left(\max_{0\le m\le \frac{3}{4}n\log n}\left|\Pi_m(j)-j\right|\le \alpha\sqrt{n\log n}\,\middle|\, j\in A^{\left\lfloor \frac{3}{4}n\log n\right\rfloor}\right)\\
&\le \min_{j\in D^n} P_{id}^n\left(R_{j,t_n}\middle| j\in A^{\left\lfloor \frac{3}{4}n\log n\right\rfloor}\right)\\
&= \min_{j\in D^n} P_{id}^n\left(R_{j,t_n}\middle| j\in A^{t_n}\right),
\end{aligned}$$

Lemma 3.2 yields

$$\begin{aligned}
&P_{id}^n\left(\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\le r\cdot E_{id}^n\left\{\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\right\}\right)\\
&\qquad\le \frac{K_{t_n}+\left(1-\mathcal{S}^2(n,\alpha)\right)K_{t_n}^2}{(1-r)^2\,\mathcal{S}^2(n,\alpha)\,K_{t_n}^2},
\end{aligned}$$

where $K_{t_n}\triangleq E_{id}^n\left\{\left|D^n\cap A^{t_n}\right|\right\}$.

A simple calculation shows that $\lim_{n\to\infty}K_{t_n}=\infty$. Theorem 2.1 (with $t_n=\left\lfloor\frac{3}{4}n\log n\right\rfloor$) implies that

$$\liminf_{n\to\infty}\mathcal{S}(n,\alpha)\ge 1-4\Psi\left(\alpha\sqrt{\frac{8}{3}}\right)=v(\alpha)>0. \tag{3.1}$$

Therefore

$$\begin{aligned}
&\limsup_{n\to\infty}P_{id}^n\left(\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\le r\cdot E_{id}^n\left\{\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\right\}\right)\\
&\qquad\le \frac{1}{(1-r)^2}\limsup_{n\to\infty}\frac{\left(1-\mathcal{S}^2(n,\alpha)\right)}{\mathcal{S}^2(n,\alpha)}\\
&\qquad\le (1-r)^{-2}\left(v^{-2}(\alpha)-1\right).
\end{aligned} \tag{3.2}$$

Note that by (3.1), for any $\delta\in(0,1)$ and sufficiently large n,

$$\begin{aligned}
E_{id}^n\left\{\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\right\}&=\sum_{j\in D^n}P_{id}^n\left(R_{j,t_n}\middle| j\in A^{t_n}\right)P_{id}^n\left(j\in A^{t_n}\right)\\
&\ge \left|D^n\right|p_{t,n}\mathcal{S}(n,\alpha)\ge \delta v(\alpha)\,\varepsilon n p_{t_n,n}.
\end{aligned}$$

Together with (3.2), this implies

$$\begin{aligned}
&\limsup_{n\to\infty}P_{id}^n\left(\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|\le r v(\alpha)\,\varepsilon n p_{t_n,n}\right)\\
&\qquad\le (1-r/\delta)^{-2}\left(v^{-2}(\alpha)-1\right).
\end{aligned}$$

By letting $\delta\to 1$, the lemma follows. ∎

Lastly, we consider $\triangle_{\alpha}(\Pi_{t_n}) \setminus A^{t_n}$. For this we will need the following proposition whose proof is given in Section 5.

Proposition 3.1. *Denote the chain corresponding to random-to-random insertions shuffle starting at $id \in S_n$ by $\left\{\Pi_m^{id}\right\}_{m=0}^{\infty}$ and the chain corresponding to random-to-random insertions shuffle with initial distribution U^n by $\left\{\Pi_m^{U^n}\right\}_{m=0}^{\infty}$. For any $n \geq 1$, $t \leq \left\lfloor \frac{3}{4} n \log n \right\rfloor$ and different $i,j \in [n]$, the two chains can be coupled so that i and j are chosen for removal at the same times for both and, under the corresponding measure $P_{id,U^n}^{i,j}$,*

$$P_{id,U^n}^{i,j}\left(\left|\Pi_t^{id}(k) - \Pi_t^{U^n}(k)\right| \geq \rho_1 \log^2 n \,\middle|\, B_{i,j}^t\right) \leq \rho_2 n^{-2},$$

for $k = i,j$, where $B_{i,j}^t$ is the event that i and j have been removed up to time t and ρ_1, ρ_2 are constants independent of n.

With the proposition, as before, we can evaluate first and second moments and apply Chebyshev's inequality.

Lemma 3.4. *Let $t_n \leq \left\lfloor \frac{3}{4} n \log n \right\rfloor$ be a sequence and let $k > 0$. Let ρ_1 be the constatnt from Proposition 3.1. Then,*

$$\limsup_{n\to\infty} P_{id}^n\left(\left|\left|\triangle_{\alpha}(\Pi_{t_n}) \setminus A^{t_n}\right| - 2\varepsilon\alpha\left(1 - p_{t_n,n}\right)\sqrt{n\log n}\right| \dots \right.$$
$$\left. \dots \geq k \cdot \sqrt{8\rho_1\alpha}\varepsilon n^{1/4} \log^{5/4} n\right) \leq \frac{1}{k^2}.$$

Proof. Denote the first time a card $j \in [n]$ is removed by $\tau(j)$. By definition, $P_{id}^n(\Pi_k(j) \in \cdot \mid \tau(j) = k) = u^n$, where u^n denotes the uniform measure on $[n]$. Since the shuffles at times $k+1, k+2, \dots$ are independent of the shuffle at times $1, \dots, k$, they are also independent of the event $\{\tau(j) = k\}$. As a marginal distribution of U^n, the stationary distribution of the chain $\{\Pi_m(j)\}_{m=0}^{\infty}$ also coincides with u^n. Therefore, for $m \geq 0$, $P_{id}^n(\Pi_{k+m}(j) \in \cdot \mid \tau(j) = k) = u^n$, and, as a mixture of measures of this form, $P_{id}^n\left(\Pi_t(j) \in \cdot \mid j \notin A^t\right)$ coincides with u^n as well.

Thus, if $\varepsilon n \geq \alpha\sqrt{n \log n}$,

$$\begin{aligned} E_{id}^n\left\{\left|\triangle_{\alpha}(\Pi_t) \setminus A^t\right|\right\} &= \sum_{j \in D^n} P_{id}^n\left(R_{j,t}^{A^c} \,\middle|\, j \notin A^t\right) P_{id}^n\left(j \notin A^t\right) \\ &= |D^n| \frac{1 + 2\left\lfloor \alpha\sqrt{n\log n}\right\rfloor}{n}\left(1 - p_{t,n}\right). \end{aligned}$$

For the second moment write

$$E_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t}\right|^{2}\right\}=\sum_{j\in D^{n}}P_{id}^{n}\left(R_{j,t}^{A^{c}}\right)+\sum_{i,j\in D:i\neq j}P_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\right).$$
$$=E_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t}\right|\right\}+\sum_{i,j\in D:i\neq j}P_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\right).$$

According to Proposition 3.1, for $i\neq j$, $n\geq 1$ and $t\leq\left\lfloor\frac{3}{4}n\log n\right\rfloor$,

$$\begin{aligned}
&P_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\middle| i,j\notin A^{t}\right)=P_{id,U^{n}}^{i,j}\left(i,j\in\triangle_{\alpha}\left(\Pi_{t}^{id}\right)\middle| B_{i,j}^{t}\right)\\
&\quad\leq P_{id,U^{n}}^{i,j}\left(\left|\Pi_{t}^{U^{n}}(j)-j\right|,\left|\Pi_{t}^{U^{n}}(i)-i\right|\leq\alpha\sqrt{n\log n}+\rho_{1}\log^{2}n\middle| B_{i,j}^{t}\right)\\
&\qquad+P_{id,U^{n}}^{i,j}\left(\left|\Pi_{t}^{id}(j)-\Pi_{t}^{U^{n}}(j)\right|\geq\rho_{1}\log^{2}n\middle| B_{i,j}^{t}\right)\\
&\qquad+P_{id,U^{n}}^{i,j}\left(\left|\Pi_{t}^{id}(i)-\Pi_{t}^{U^{n}}(i)\right|\geq\rho_{1}\log^{2}n\middle| B_{i,j}^{t}\right)\\
&\quad\leq\frac{\left(1+2\left\lfloor\alpha\sqrt{n\log n}+\rho_{1}\log^{2}n\right\rfloor\right)^{2}}{n\left(n-1\right)}+2\rho_{2}n^{-2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_{i,j\in D:i\neq j}P_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\right)=\sum_{i,j\in D:i\neq j}P_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\middle| i,j\notin A^{t}\right)P_{id}^{n}\left(i,j\notin A^{t}\right)\\
&\quad\leq\left|D^{n}\right|^{2}\left(\frac{\left(1+2\left\lfloor\alpha\sqrt{n\log n}+\rho_{1}\log^{2}n\right\rfloor\right)^{2}}{n\left(n-1\right)}+2\rho_{2}n^{-2}\right)\\
&\quad\times\left(1-2p_{t,n}+\left(\frac{n-2}{n}\right)^{t}\right).
\end{aligned}$$

By straightforward algebra, the above yield

$$Var_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t_{n}}\right)\setminus A^{t}\right|\right\}\leq 8\rho_{1}\alpha\varepsilon^{2}n^{1/2}\log^{5/2}n+O\left(\sqrt{n\log n}\right).$$

Applying Chebyshev's inequality and taking the limit completes the proof.

□

Remark 3.1. Based on Lemma 3.1 and Lemma 3.4, combined with a 'coupon-collecting' argument to bound the probability $P_{id}^{n}\left(\left|A^{t_{n}}\right|/\sqrt{n\log n}>\delta\right)$, it is easy to see that for t_n as in Theorem 1.1, as long as c_n does not approach ∞ too fast,

$$\left|\triangle_{\alpha}\right|/\left(2\varepsilon\alpha\sqrt{n\log n}\right)\to 1\quad\text{in probability},$$

under U^n and P_{id}^{n}.

4. Proof of Theorem 1.1. In order to distinguish the two measures we consider the deviation of $|\triangle_\alpha|$ from $2\varepsilon\alpha\sqrt{n\log n}$. Let $k,c>0$, and let α such that $v(\alpha)>0$ (where $v(\alpha)$ was defined in Lemma 3.3) and assume t_n is as in the theorem. Suppose that for some n

$$\left|\left|\triangle_\alpha(\Pi_t)\setminus A^{t_n}\right|-2\varepsilon\alpha(1-p_{t_n,n})\sqrt{n\log n}\right|<k\sqrt{8\rho_1\alpha}\varepsilon n^{1/4}\log^{5/4}n, \tag{4.1}$$

and

$$\left|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}\right|>\frac{1}{2}v(\alpha)\varepsilon n p_{t_n,n}. \tag{4.2}$$

Then, if n is sufficiently large,

$$\begin{aligned}
&|\triangle_\alpha(\Pi_{t_n})|-2\varepsilon\alpha\sqrt{n\log n}\\
&\geq \varepsilon n p_{t_n,n}\left(\frac{1}{2}v(\alpha)-2\alpha\sqrt{n\log n}/n\right)-k\sqrt{8\rho_1\alpha}\varepsilon n^{1/4}\log^{5/4}n\\
&=\frac{1}{2}\varepsilon n p_{t_n,n}(v(\alpha)-o(1))-k\sqrt{8\rho_1\alpha}\varepsilon n^{1/4}\log^{5/4}n\\
&\geq cn^{1/4}\log^{5/4}n,
\end{aligned} \tag{4.3}$$

where the last inequality follows from the following calculation.

Writing

$$\log\frac{np_{t_n,n}}{n^{1/4}\log^{5/4}n}=\frac{3}{4}\log n-\frac{5}{4}\log\log n+\log p_{t_n,n},$$

substituting $p_{t_n,n}$ and t_n and using the fact that $\log(1+x)=x+O(x^2)$ as $x\to 0$ we arrive at

$$\log\frac{np_{t_n,n}}{n^{1/4}\log^{5/4}n}=c_n+o(1)\to\infty.$$

Now, since, for large n, (4.1) and (4.2) imply (4.3), by a union bound, Lemmas 3.3 and 3.4 imply

$$\begin{aligned}
&\liminf_{n\to\infty}P_{id}^n\left(|\triangle_\alpha(\Pi_{t_n})|-2\varepsilon\alpha\sqrt{n\log n}\geq cn^{1/4}\log^{5/4}n\right)\\
&\geq 1-\frac{1}{k^2}-\left(1-\frac{1}{2}\right)^{-2}\left(v^{-2}(\alpha)-1\right)\xrightarrow{k,\alpha\to\infty}1.
\end{aligned}$$

In addition, from Lemma 3.1,

$$\limsup_{n\to\infty}U^n\left(|\triangle_\alpha(\sigma)|-2\varepsilon\alpha\sqrt{n\log n}\geq cn^{1/4}\log^{5/4}n\right)\leq\frac{1}{k^2}\xrightarrow{k\to\infty}0.$$

Thus, since k and α were arbitrary,

$$\liminf_{n\to\infty}\|P_{id}^n(\Pi_t\in\cdot)-U^n\|_{TV}=0.$$

$\square$

5. Proofs of Lemma 3.2 and Proposition 3.1. In this section we prove Lemma 3.2 and Proposition 3.1. We begin with the lemma, which, since cards are chosen independently each step, is basically a result related to the Coupon-Collector's problem (see Feller (1968)).

5.1. *Proof of Lemma 3.2.* By our assumption,

$$E_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|\right\}=\sum_{j\in D}P_{id}^{n}\left(j\in B|\,j\in A^{t}\right)P_{id}^{n}\left(j\in A^{t}\right)\geq cK.$$

Write

$$\begin{aligned}
&E_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|^{2}\right\}\\
&\quad\leq E_{id}^{n}\left\{\left|D\cap A^{t}\right|^{2}\right\}=\sum_{i,j\in D}P_{id}^{n}\left(i,j\in A^{t}\right)\\
&\quad=\sum_{j\in D}P_{id}^{n}\left(j\in A^{t}\right)+\sum_{i,j\in D:i\neq j}P_{id}^{n}\left(i,j\in A^{t}\right)\\
&\quad=K+|D|\left(|D|-1\right)\left(\frac{n-2}{n}\right)^{t}.
\end{aligned}$$

Since $K=|D|((n-1)/n)^{t}$ it follows that

$$E_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|^{2}\right\}\leq K+K^{2},$$

therefore

$$Var_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|\right\}\leq K+\left(1-c^{2}\right)K^{2}.$$

Applying Chebyshev's inequality completes the proof. □

5.2. *Proof of Proposition 3.1.* Define a coupling of Π_{m}^{id} and $\Pi_{m}^{U^{n}}$ as follows. At each step remove a card from each of the decks as described below, choose a random position and reinsert both cards in this position in both decks. Denote the positions of i and j in both decks at time m (i.e. after the m-th shuffle) by

$$\Lambda_{m}\triangleq\left\{\Pi_{m}^{id}\left(i\right),\Pi_{m}^{id}\left(j\right),\Pi_{m}^{U^{n}}\left(i\right),\Pi_{m}^{U^{n}}\left(j\right)\right\}.$$

On the m-th shuffle, with probability $1-\left|\Lambda_{m-1}\right|/n$ choose a random *position* from $[n]\setminus\Lambda_{m-1}$, and remove from both decks the card in this position. With probability $1/n$ remove the card *numbered* i from both decks. With

probability $1/n$ do the same for j. Lastly, with probability $(|\Lambda_{m-1}|-2)/n$ choose a random position from

$$\Lambda_{m-1}\setminus\left\{\Pi_{m-1}^{id}(i),\Pi_{m-1}^{id}(j)\right\}$$

and a random position from

$$\Lambda_{m-1}\setminus\left\{\Pi_{m-1}^{U^n}(i),\Pi_{m-1}^{U^n}(j)\right\}$$

and remove the cards in these positions from the decks respectively. Denote the measure corresponding to the coupling by $P_{id,U^n}^{i,j}$.

Note that by the definition of the coupling, $P_{id,U^n}^{i,j}=P_{id,U^n}^{j,i}$. Therefore it is enough to prove the inequality in the proposition only for j. Also note that we can, and shall, assume without loss of generality that $n>9$, since if the inequality is true for such n, by adjusting the constants ρ_1, ρ_2 it will hold for smaller n.

Let $x\geq 1$, let $t_0=\left\lfloor\frac{3}{4}n\log n\right\rfloor$ and fix $t\leq t_0$. Define for each $m\geq 1$,

$$\Upsilon_m=\left[\min\left\{\Pi_m^{id}(j),\Pi_m^{U^n}(j)\right\},\max\left\{\Pi_m^{id}(j),\Pi_m^{U^n}(j)\right\}\right]\cap[n],$$

$$I_m=|\Upsilon_m\setminus\Lambda_m|\quad\text{and}\quad O_m=|[n]\setminus(\Upsilon_m\cup\Lambda_m)|.$$

That is, I_m is the number of positions *in* the range determined by the positions of j in both decks and O_m is the number of positions *out* of this range, with both sets in addition excluding Λ_m.

Let H_m denote the event that the card which was removed on the m-th shuffle was in a position in $[n]\setminus\Lambda_{m-1}$ and was not reinserted in a position which is greater by 1 from a position in Λ_m and not reinserted in position 1. Define $\eta(j)$ to be the last time j is removed before or at time t, and $\eta(j)=0$, if it has not been removed up to t. Define

$$\vartheta_x=\min\left\{m>\eta(j):\left|\Pi_m^{id}(j)-\Pi_m^{U^n}(j)\right|\geq x\right\}.$$

Finally, define the process

$$L_m=\sum_{k=\eta(j)+1}^{\min\{\eta(j)+m,\vartheta_x\}}(I_k-I_{k-1})\,\mathbf{1}_{H_k},$$

where $\mathbf{1}_F$ is the indicator function of a set F. The reason for defining L_m is that, as we shall see, conditioned on $B_{i,j}^t$, its increments satisfy the condition of Proposition 2.1.

Assume $B_{i,j}^t$ occurs. If on the m-th shuffle the card numbered j is chosen, then $\Pi_m^{id}(j) = \Pi_m^{U^n}(j)$, and thus $I_{\eta(j)} = 0$. Hence, for any time $\eta(j) < s \le \min\{t, \vartheta_x\}$,

$$\begin{aligned} I_s &= \sum_{m=\eta(j)+1}^{s} (I_m - I_{m-1}) \mathbf{1}_{H_m} + \sum_{m=\eta(j)+1}^{s} (I_m - I_{m-1}) \mathbf{1}_{H_m^c} \\ &= L_{s-\eta(j)} + \sum_{m=\eta(j)+1}^{s} (I_m - I_{m-1}) \mathbf{1}_{H_m^c}. \end{aligned}$$

For any $\eta(j) < m \le t$, j is not removed at time m, and therefore, as can be easily seen, $\left|(I_m - I_{m-1}) \mathbf{1}_{H_m^c}\right| \le 2$.

Now, still under the assumption that $B_{i,j}^t$ occurs, if $\vartheta_x \le t$, then, since by definition

$$0 \le \left|\Pi_t^{id}(j) - \Pi_t^{U^n}(j)\right| - I_t \le 3,$$

it holds that

$$\max_{0\le m\le t} L_m + 2\sum_{m=1}^{t} \mathbf{1}_{H_m^c} \ge L_{\vartheta_x - \eta(j)} + \sum_{m=\eta(j)+1}^{s} (I_m - I_{m-1}) \mathbf{1}_{H_m^c} = I_{\vartheta_x} \ge x - 3.$$

Therefore for $0 \le u$, δ such that $u + \delta \le x - 3$, a union bound gives

$$\begin{aligned} &P_{id,U^n}^{i,j}\left(\left|\Pi_t^{id}(j) - \Pi_t^{U^n}(j)\right| \ge x \,\middle|\, B_{i,j}^t\right) \le P_{id,U^n}^{i,j}\left(\vartheta_x \le t \middle| B_{i,j}^t\right) \\ &\le P_{id,U^n}^{i,j}\left(\max_{0\le m\le t} L_m \ge u \,\middle|\, B_{i,j}^t\right) + P_{id,U^n}^{i,j}\left(\sum_{m=1}^{t} \mathbf{1}_{H_m^c} \ge \delta/2 \,\middle|\, B_{i,j}^t\right). \end{aligned} \tag{5.1}$$

Next, we apply Proposition 2.1 to L_m conditioned on $B_{i,j}^t$, in order to bound its exceedance probability in (5.1). For brevity, denote $m' = m + \eta(j)$. First of all, note that given that $m' + 1 \le \vartheta_x$, $L_{m+1} = L_m + 1$ (similarly, $L_{m+1} = L_m - 1$) if and only if the position of the card that is chosen for removal on step $m'+1$ is in $[n] \setminus (\Upsilon_{m'} \cup \Lambda_{m'})$ (respectively, $\Upsilon_{m'} \setminus \Lambda_{m'}$), and its position after reinsertion is greater by 1 from some position in $\Upsilon_{m'+1} \setminus \Lambda_{m'+1}$ ($[n] \setminus (\Upsilon_{m'+1} \cup \Lambda_{m'+1})$). Therefore, for $d = \pm 1$, the number of different insertion shuffles (i.e., choices of a card and a position) such that $L_{m+1} = L_m + d$ is $I_{m'} O_{m'}$. Similarly, it can be easily seen that the number of different insertion shuffles such that $H_{m'+1}$ occurs is $(I_{m'} + O_{m'})(I_{m'} + O_{m'} - 1)$.

Thus, for $d = \pm 1$,

$$\begin{aligned}
&P^{i,j}_{id,U^n}\left(L_{m+1} = k_m + d \middle| B^t_{i,j},\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m, \dots \right. \\
&\qquad\qquad \left. H_{m'+1},\, m'+1 \le \vartheta_x,\, I_{m'} = a,\, O_{m'} = b\right) \\
&\quad = P^{i,j}_{id,U^n}\left(I_{m'+1} - I_{m'} = d \middle| I_{m'} = a,\, O_{m'} = b,\, H_{m'+1}\right) \\
&\quad = \begin{cases} \frac{ab}{(a+b)(a+b-1)} & \text{if } a+b > 1, \\ 0 & \text{if } a+b \le 1. \end{cases}
\end{aligned} \tag{5.2}$$

Since the only nonzero increments of L_m are ± 1 and $\{L_{m+1} \neq L_m\} \subset \{H_{m'+1},\, m'+1 \le \vartheta_x\}$, it follows that

$$\begin{aligned}
&P^{i,j}_{id,U^n}\left(L_{m+1} = k_m + 1 \middle| B^t_{i,j},\, L_{m+1} \neq L_m,\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\right) \\
&\quad = P^{i,j}_{id,U^n}\left(L_{m+1} = k_m - 1 \middle| B^t_{i,j}\,, L_{m+1} \neq L_m,\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\right) = \frac{1}{2}.
\end{aligned}$$

Additionally, if

$$P^{i,j}_{id,U^n}\left(B^t_{i,j},\, L_{m+1} \neq L_m,\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\right) > 0,$$

then

$$\begin{aligned}
&P^{i,j}_{id,U^n}\left(L_{m+1} \neq L_m \middle| B^t_{i,j},\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\right) \\
&\quad \le P^{i,j}_{id,U^n}\left(L_{m+1} \neq L_m \middle| B^t_{i,j},\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\,,\, H_{m'+1},\, m'+1 \le \vartheta_x\right) \\
&\quad \le \max_{a,b} P^{i,j}_{id,U^n}\left(L_{m+1} \neq L_m \middle| B^t_{i,j},\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\,, \dots \right. \\
&\qquad \left. H_{m'+1},\, m'+1 \le \vartheta_x,\, I_{m'} = a,\, O_{m'} = b\right),
\end{aligned}$$

where the maximum is taken over all values of a and b such that the conditional probability is defined. By definition, if $m'+1 \le \vartheta_x$, then $I_{m'} \le x$. In addition, for any m, $O_m + I_m \ge n-4$. Hence, from the last inequality and (5.2) we have (recall we assumed that $n > 9$),

$$\begin{aligned}
&P^{i,j}_{id,U^n}\left(L_{m+1} \neq L_m \middle| B^t_{i,j},\, \{L_p\}_{p=1}^m = \{k_p\}_{p=1}^m\right) \\
&\qquad \le \max_{\substack{0 \le a \le x \\ a+b \ge n-4}} \frac{2ab}{(a+b)\,(a+b-1)} \le \frac{2x}{n-4}.
\end{aligned}$$

By Proposition 2.1 applied to L_m as a process conditioned on $B_{i,j}^t$, we have, for $0 \leq u \leq x$,

$$P_{id,U^n}^{i,j}\left(\max_{0\leq m\leq t} L_m \geq u \middle| B_{i,j}^t\right) \leq P\left(\max_{0\leq m\leq t} S_m^L \geq u\right), \tag{5.3}$$

where S_m^L is a random walk with increment distribution

$$\mu^L(1) = \mu^L(-1) = \frac{x}{n-4},\ \mu^L(0) = 1 - \frac{2x}{n-4}.$$

Recall that $t \leq t_0 = \left\lfloor \frac{3}{4} n \log n \right\rfloor$. By Lévy's and Bernstein's inequalities (Petrov (1995), Theorem 2.8), for $0 \leq u \leq x$,

$$\begin{aligned} &P\left(\max_{0\leq m\leq t} S_m^L \geq u\right) \leq P\left(\max_{0\leq m\leq t_0} S_m^L \geq u\right) \\ &\leq 2P\left(S_{t_0}^L \geq u\right) \leq 2\exp\left\{-\frac{u^2(n-4)}{x8t_0}\right\} \leq 2\exp\left\{-\frac{u^2}{12x\log n}\right\}. \end{aligned} \tag{5.4}$$

Now let us bound the other summand of (5.1). For $n > 9$, since $|\Lambda_m| \leq 4$, it can be easily seen that

$$\begin{aligned} &P_{id,U^n}^{i,j}\left(\sum_{m=1}^{t} \mathbf{1}_{H_m^c} \geq \frac{\delta}{2} \middle| B_{i,j}^t\right) \\ &\leq P\left(Q_t \geq \frac{\delta}{2} - 2\right) \leq P\left(Q_{t_0} \geq \frac{\delta}{2} - 2\right), \end{aligned} \tag{5.5}$$

where $Q_t \sim \mathrm{Bin}\left(t, \frac{9}{n}\right)$ and similarly for t_0. For $\delta \geq 27\log n + 4$, Bernstein's inequality gives

$$\begin{aligned} P\left(Q_{t_0} \geq \frac{\delta}{2} - 2\right) &\leq \exp\left\{-\frac{\delta/2 - 2 - 9t_0/n}{4}\right\} \\ &\leq \exp\left\{-\frac{\delta}{8} + \frac{1}{2} + \frac{27}{16}\log n\right\}. \end{aligned} \tag{5.6}$$

From (5.1) and (5.3)-(5.6), choosing $x = 100\log^2 n$, $u = 50\log^2 n$ and $\delta = 30\log n + 4$, for example, we obtain

$$P_{id,U^n}\left(\left|\Pi_t^{id}(j) - \Pi_t^{U^n}(j)\right| \geq 100\log^2 n \middle| B_{i,j}^t\right) \leq 3n^{-2}.$$

$\square$

Acknowledgements. This work arose from a graduate course taught by Professor Ross G. Pinsky. I am grateful to him and my advisor, Professor Robert J. Adler, for valuable comments on earlier versions of this paper.

References.

Diaconis, P. (2003). Mathematical developments from the analysis of riffle shuffling. In *Groups, combinatorics & geometry (Durham, 2001)* 73–97. World Sci. Publ., River Edge, NJ.

Diaconis, P. and Saloff-Coste, L. (1993). Comparison techniques for random walk on finite groups. *Ann. Probab.* **21** 2131–2156.

Diaconis, P. and Saloff-Coste, L. (1995). Random walks on finite groups: a survey of analytic techniques. In *Probability measures on groups and related structures, XI (Oberwolfach, 1994)* 44–75. World Sci. Publ., River Edge, NJ.

Feller, W. (1968). *An introduction to probability theory and its applications. Vol. I. Third edition*. John Wiley & Sons Inc., New York.

Petrov, V. V. (1995). *Limit theorems of probability theory. Oxford Studies in Probability* **4**. The Clarendon Press Oxford University Press, New York.

Saloff-Coste, L. and Zúñiga, J. (2008). Refined estimates for some basic random walks on the symmetric and alternating groups. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 359–392.

Uyemura-Reyes, J.-C. (2002). Random walk, semi-direct products, and card shuffling PhD thesis, Stanford University.

Eliran Subag
Electrical Engineering
Technion, Haifa, Israel 32000
E-mail: elirans@techunix.technion.ac.il